E C O N O M I C S B U L L E T I N

Latent Leading and Coincident Factors Model with Markov−Switching Dynamics

Konstantin Kholodilin
Institut des recherches économiques et sociales

Abstract

This paper introduces a two−factor model of leading and coincident economic indicators. The common leading factor is assumed to Granger−cause the common coincident factor. This property is used to estimate the two common factors simultaneously and hence more efficiently. Two models of the latent leading and coincident factors are studied: a model with linear dynamics and a model with Markov−switching dynamics introduced through the leading factor intercept term. The first model encompasses the comovements between the individual time series. The second model, moreover, takes care of possible asymmetries between the business cycle regimes.

Citation: Kholodilin, Konstantin, (2001) "Latent Leading and Coincident Factors Model with Markov−Switching Dynamics." *Economics Bulletin,* Vol. 3, No. 7 pp. 1−13
Submitted: June 25, 2001. **Accepted:** July 17, 2001.
URL: http://www.economicsbulletin.com/2001/volume3/EB−01C50001A.pdf

1 Introduction

In the modern macroeconomic literature many efforts are devoted to identifying a hypothetical coincident economic indicator which represents a general economic activity and allows to trace the evolution of the business cycle. It is designed to serve as a reference time series to judge about the state of the affairs in the economy. The most prominent examples of the one-factor models with the linear dynamics is Stock and Watson (1988), while with the Markov-switching dynamics these are Chauvet (1998), Kim and Yoo (1995).

With respect to this common coincident indicator one can then define the leading and lagging macroeconomic variables. The former of these series are especially important since they permit to predict the changes in the state of the economy before they have occurred.

Normally, however, the leading series are not aggregated into a common leading factor. The evolution of the common coincident factor is conditioned on each of them individually, either directly through a VAR system of the common coincident factor and individual leading observed time series as in Stock and Watson (1988), Chauvet and Potter (2000) or via the time-varying transition probabilities which depend on the individual leading variables as in Kim and Yoo (1995).

This paper introduces a two-factor model where one of the latent factors is postulated as a common leading indicator, while the second factor is taken to be the common coincident indicator. There assumed to exist a one-way Granger causality coming from the former common factor to the latter one. The common leading and coincident factors are estimated from a set of the observed time series which is split into a subset of leading and a subset of coincident variables.

First, we consider a linear model with leading and coincident factor following an AR process. Next, we add a regime-switching dynamics to take care of the possible asymmetries between the recession and expansion phases of the business cycle captured by both common latent factors.

The linear specification of the two-factor model is presented in the section two, while section three contains a description of the model with nonlinear dynamics. In the section four we apply our models to the artificial data in order to see how well these models reflect the true data-generating process. Section five concludes the paper. All the tables and graphs are put into the Appendix following the list of references.

2 Linear model

We consider a set of the observed time series, some of which may be defined as leading while the rest of them are treated as the coincident series. The common dynamics of the time series belonging to each of these groups are underlined by a common factor: leading corresponding to the first group and coincident corresponding to the second group. The idiosyncratic dynamics of each time series in particular are captured by one specific factor per each observed time series. Therefore the model can be written as follows:

$$\Delta y_t = \Gamma \Delta f_t + u_t \tag{1}$$

where $\Delta y_t = (\Delta y_{Lt} \mid \Delta y_{Ct})'$ is the $n \times 1$ vector of the observed time series in the first differences; $\Delta f_t = (\Delta f_{Lt} \mid \Delta f_{Ct})'$ is the 2×1 vector of the latent common factors in the first differences; $u_t = (u_{Lt} \mid u_{Ct})'$ is the $n \times 1$ vector of the latent specific factors; Γ is the $n \times 2$ factor loadings matrix linking the observed series with the common factors.

The dynamics of the latent common factors can be described in terms of a VAR model:

$$\Delta f_t = \mu + \Phi(L) \Delta f_{t-1} + \varepsilon_t \tag{2}$$

where μ is the 2×1 vector of the constant intercepts; $\Phi(L)$ is the sequence of p ($p = \max\{p_L, p_C\}$, where p_L is the order of the AR polynomial of the leading factor, and p_C is the order of the AR polynomial of the coincident factor) 2×2 lag polynomial matrices; ε_t is the 2×1 vector of the serially and mutually uncorrelated common factor disturbances:

$$\varepsilon_t \sim NID\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_L^2 & 0 \\ 0 & \sigma_C^2 \end{pmatrix} \right)$$

We assume that the leading factor Granger-causes the coincident factor but not vice versa. This assumption means that the matrices Φ_i $(i = 1, ..., p)$ are diagonal or lower diagonal for all i. For simplicity we suppose that the causality from the leading to the coincident factor is transmitted only at one lag, say τ. Thus, if $i \neq \tau$,

$$\Phi_i = \begin{pmatrix} \phi_{L,i} & 0 \\ 0 & \phi_{C,i} \end{pmatrix}$$

and if $i = \tau$,

$$\Phi_i = \begin{pmatrix} \phi_{L,i} & 0 \\ \phi_{CL,i} & \phi_{C,i} \end{pmatrix}$$

The idiosyncratic factors are by definition mutually independent and are modelled as the AR processes:

$$u_t = \Psi(L) u_{t-1} + \eta_t \tag{3}$$

where $\Psi(L)$ is the sequence of q ($q = \max\{q_{1,...}, q_n\}$, where q_i is the order of the AR polynomial of the $i-th$ idiosyncratic factor) $n \times n$ diagonal lag polynomial matrices and η_t is the $n \times 1$ vector of the mutually and serially uncorrelated Gaussian shocks:

$$\eta_t \sim \left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix} \right)$$

To estimate this model we express it in a state-space form:

Measurement equation:

$$\Delta y_t = A\beta_t \tag{4}$$

Transition equation:

$$\beta_t = \alpha + C\beta_{t-1} + v_t \tag{5}$$

where $\beta_t = (f_t|u_t)'$ is the state vector containing stacked on top of each other vector of common factors and the vector of specific factors; v_t is the vector of the common and idiosyncratic factors' disturbances with mean zero and variance-covariance matrix Q; α is the vector of intercepts.

$A = \begin{pmatrix} \gamma_L & O_{n_L \times (r+p_C)} & i_{q_1} & ... & 0 \\ O_{n_L \times r} & \gamma_C & 0 & ... & i_{q_n} \end{pmatrix}$

where γ_L is the $n_L \times 1$ vector of the leading factor loadings; $O_{n \times m}$ is $n \times m$ matrix of zeros; i_m is the first row of the $m \times m$ identity matrix, and $r = \max\{p_L, \tau\}$.

$C = \begin{pmatrix} \Phi^L & & & & 0 \\ \Phi^{CL} & \Phi^C & & & \\ & & \Psi^1 & & \\ & & & \ddots & \\ 0 & & & & \Psi^n \end{pmatrix}$

where Φ^L is the $r \times r$ matrix:

$\Phi^L = \begin{pmatrix} \phi_L & o'_{r-p_L} \\ I_{r-1} & O_{(r-1)\times(r-p_L)} \end{pmatrix}$

where ϕ_L is the $1 \times p_L$ row vector of the AR coefficients of the leading factor, I_n is the $n \times n$ identity matrix, and o_m is the $m \times 1$ vector of zeros.

$\Phi^C = \begin{pmatrix} \phi_C & 0 \\ I_{p_C-1} & o_{p_C-1} \end{pmatrix}$

The matrices $\Psi^1, ..., \Psi^n$ have the same structure as Φ^C.

$\Phi^{CL} = \begin{pmatrix} o'_r \\ \phi_{CL} \end{pmatrix}$

where ϕ_{CL} is the $1 \times r$ vector of zeros with ϕ_{CL}, τ at the $\tau - th$ position.

The unknown parameters and the latent factors may be estimated using Kalman filter recursions. To save space we will not present them here, referring the reader, for instance, to Hamilton (1994) who gives very clear and systematic explanation of the Kalman filter methodology.

3 Nonlinear model

It was observed by many authors, among them by Diebold and Rudebusch (1996) that the model of the business cycle would be incomplete if it would not take into account both the comovement of various macroeconomic variables and the asymmetries between the phases of the cycle. The linear model presented in the previous section incorporates the phenomenon of the simultaneous changes in the levels of different individual time series. However, it lacks a mechanism which would reflect the qualitatively different behavior of these series during recessions and expansions. One of the ways to introduce this mechanism in our model is to add to it the regime-switching dynamics.

The Markov-switching dynamics is introduced through the leading factor intercept:

$$\Delta f_t = \mu(s_t) + \Phi(L)\Delta f_{t-1} + \varepsilon_t \tag{6}$$

where $\mu(s_t) = (\mu_L(s_t), ..., 0)'$.

s_t is the unobserved regime variable. In the two-regime (expansion-recession) case it takes two values: 0 or 1. Depending on the regime, the leading factor intercept assumes different values: low in recessions and high in expansions. Thus, the common factors grow faster during the upswings and slower (or even have negative growth rate) during the downswings of the economy.

The changes in the regimes are governed by the first-order Markov chain process, which is summarized by the transition probabilities matrix:

	$s_{t-1} = 0$	$s_{t-1} = 1$
$s_t = 0$	p_{11}	$1 - p_{22}$
$s_t = 1$	$1 - p_{11}$	p_{22}

where $p_{ij} = prob(s_t = j | s_{t-1} = i)$.

The rest of the equations of the model remains unchanged. The state-space representation of the nonlinear two-factor model may be written as:

Measurement equation:

$$\Delta y_t = A\beta_t \tag{7}$$

Transition equation:

$$\beta_t = \alpha(s_t) + C\beta_{t-1} + v_t \tag{8}$$

where $\alpha(s_t) = (\mu_L(s_t), ..., 0)'$.

It is worthwhile to notice that, since it is the dynamics of the common leading factor which include the state-dependent intercept in the current period, the conditional regime probabilities predicting the occurrence of recessions or expansions of the coincident factor are simply the conditional regime probabilities computed for the leading factor shifted forward for τ periods. Thus, the conditional regime probabilities estimated using the above model provide us with the τ-periods ahead forecast of the coincident factor regimes.

All the other system matrices are as in the linear model. Thus, we have a model expressed in the state-space form and having Markov-switching dynamics. Again, we will not reproduce here all the relevant recursions which are necessary to estimate the parameters and the unobserved state vector. On the estimation of the common factor models with Markov switching one can read in Kim (1994) or Kim and Nelson (1999).

4 Artificial examples

For the linear case we have generated two common latent factors and five individual observable series. The first two observed time series are leading, while the three remaining are the coincident. Both the common factors (in fact, their first differences, not levels) and the idiosyncratic components are modelled as the stationary AR(1) processes. The coincident factor is positively affected by the leading factor at the lag $\tau = 3$. The true parameters of the DGP are presented in the column two of the Table 1 of the Appendix. The length of all these series is 540 observations, which is comparable to the length of an ordinary Post World War II monthly time series for the US economy.

To identify the model, we set the factor loadings of the first observable variable in each subset - leading and coincident - equal to unity. Thus, we estimate only three of five factor loadings: one for the leading factor and two for the coincident factor. The model is estimated by the maximum likelihood. The estimated parameters together with the standard errors and the p-values are reproduced in the Table 1. The mere observation of the true and estimated parameters' values shows that the latter are sufficiently close to the former suggesting that the proposed model estimates the parameters generated process accurately enough.

Figure 1 compares the true and estimated leading and coincident factors in levels. To obtain the time series in levels we consecutively sum up their first differences setting the first observation equal to zero. Figure 1 display very high degree of similarity of the simulated and estimated common factors, especially in the case of the latent leading factor.

In the case of the Markov-switching dynamics the length of the series is also 540. The first two observable time series are leading, meanwhile the last three series are coincident. The coincident factor is again correlated to the leading factor with a lag of three periods. The same identifying normalization - by setting the factor loadings of the first observed time series in each group of the variables - is used. The parameters of the true DGP are presented in the second column of Table 2 of the Appendix. The estimates replicate the true parameters with a sufficiently high degree of precision.

The true and estimated common factors are shown on the Figure 2. Again, as in the case of the linear model, the estimated common factors series are very similar to the simulated common factors.

Figure 3 displays the filtered and smoothed conditional probabilities of the economy being in the recession, that is, in the low growth rate regime. These are compared with the true low regime which is depicted in the lower panels of the figure. Ones correspond to the downswing in the simulated economy, while the zeros stand for the upswings. The estimated model captures the recession dates pretty well. However, the smoothed recession probabilities sometimes miss the recessions when those have a very short duration. Thus, the smoothed probabilities turn out to be a more conservative dating tool than the filtered probabilities.

5 Summary

In this paper we have introduced a common dynamic factor model with two factors: leading and coincident. Each of them represents the common dynamics of a corresponding subset of the observed time series which are classified as being leading or coincident with respect to some hypothetical "state of the economy". The common leading factor Granger-causes the coincident factor, thus allowing to use the former in the predictions of the future values of the latter. This permits to improve the forecasting of the coincident factor because of the additional information coming from the leading variables.

We consider two models: a model with the linear dynamics and a model with the regime switching. The second model allows to take care of the asymmetries which may characterize different phases of the business cycle and therefore is more complete from the standpoint of the Burns and Mitchell's definition of the business cycle as interpreted by Diebold and Rudebusch (1996).

Both models are illustrated on two artificial examples, which show a high enough fitting ability of these models when they correspond to the true data-generating process.

References

[1] Chauvet M. (1998) "An Econometric Characterization of Business Cycle Dynamics with Factor Structure and Regime Switching" *International Economic Review* **39,** 969-96.

[2] Chauvet M., Potter S. (2000) "Coincident and Leading Indicators of the Stock Market" *Journal of Empirical Finance* **7**, 87-111.

[3] Diebold F.X., Rudebusch G.D. (1996) "Measuring Business Cycles: A Modern Perspective" *The Review of Economics and Statistics* **78**, 67-77.

[4] Hamilton J.D. (1994) *Time Series Analysis*. New Jersey: Princeton University Press.

[5] Kim C.-J. (1994) "Dynamic Linear Models with Markov-Switching" *Journal of Econometrics* **60**, 1-22.

[6] Kim C.-J., Nelson C.R. (1999) *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. Cambridge: MIT Press.

[7] Kim M.-J., Yoo J.-S. (1995) "New Index of Coincident Indicators: A Multivariate Markov Switching Factor Model Approach" *Journal of Monetary Economics* **36**, 607-30.

[8] Stock J.H., Watson M.W. (1988) "A Probability Model of the Coincident Economic Indicators", NBER working paper 2772.

6 Appendix

Table 1. True and estimated parameters of the linear two-factor model

Parameter	True	Estimated	St. error	p-value
γ_1	1	-	-	-
γ_2	0.9	0.91	0.03	0.0
γ_3	1	-	-	-
γ_4	2	1.99	0.03	0.0
γ_5	1.7	1.70	0.02	0.0
ϕ_L	0.8	0.80	0.03	0.0
ϕ_C	0.7	0.68	0.02	0.0
$\phi_{CL,3}$	0.5	0.53	0.04	0.0
ψ_1	-0.3	-0.30	0.06	0.0
ψ_2	-0.7	-0.67	0.04	0.0
ψ_3	-0.5	-0.54	0.04	0.0
ψ_4	-0.2	-0.19	0.06	0.0
ψ_5	-0.8	-0.77	0.03	0.0
σ_1^2	0.25	0.24	0.02	0.0
σ_2^2	0.36	0.38	0.03	0.0
σ_3^2	0.16	0.16	0.01	0.0
σ_4^2	0.49	0.50	0.05	0.0
σ_5^2	0.81	0.79	0.06	0.0
σ_L^2	0.25	0.26	0.03	0.0
σ_C^2	0.36	0.31	0.03	0.0

Table 2. True and estimated parameters of the nonlinear two-factor model

Parameter	True	Estimated	St. error	p-value
p_{11}	0.95	0.94	0.01	0.0
p_{22}	0.87	0.82	0.04	0.0
μ_{L1}	0.4	0.41	0.03	0.0
μ_{L2}	-0.6	-0.66	0.05	0.0
γ_1	1	-	-	-
γ_2	0.9	0.90	0.01	0.0
γ_3	1	-	-	-
γ_4	2	2.01	0.01	0.0
γ_5	1.7	1.70	0.01	0.0
ϕ_L	0.8	0.79	0.02	0.0
ϕ_C	0.7	0.75	0.02	0.0
$\phi_{CL,3}$	0.5	0.43	0.02	0.0
ψ_1	-0.3	-0.35	0.05	0.0
ψ_2	-0.7	-0.67	0.04	0.0
ψ_3	-0.5	-0.50	0.05	0.0
ψ_4	-0.2	-0.30	0.06	0.0
ψ_5	-0.8	-0.79	0.03	0.0
σ_1^2	0.25	0.24	0.02	0.0
σ_2^2	0.36	0.36	0.03	0.0
σ_3^2	0.16	0.17	0.01	0.0
σ_4^2	0.49	0.44	0.04	0.0
σ_5^2	0.81	0.84	0.06	0.0
σ_L^2	0.16	0.15	0.02	0.0
σ_C^2	0.36	0.34	0.03	0.0

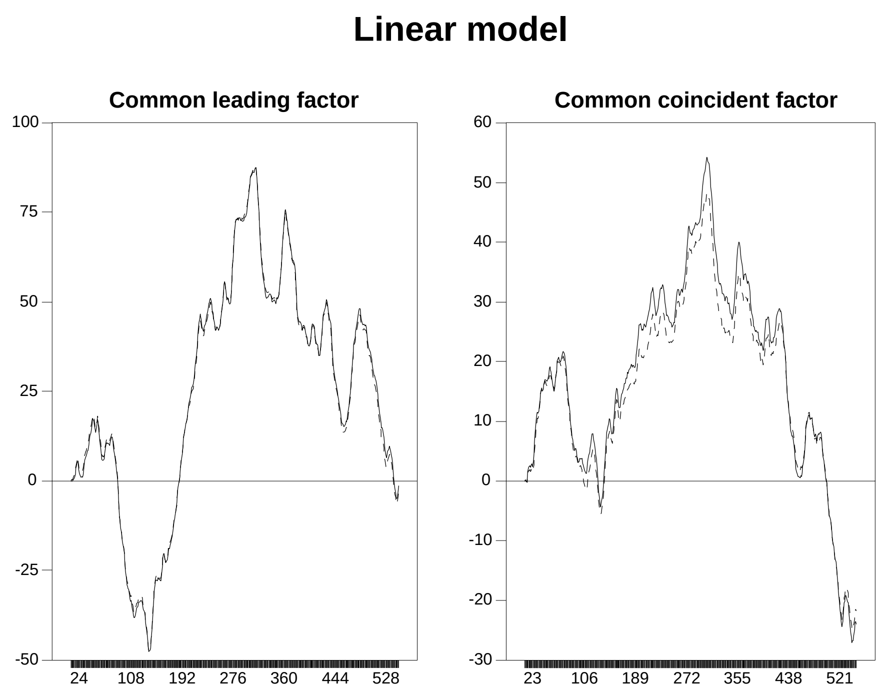

Figure 1: True and estimated common factors

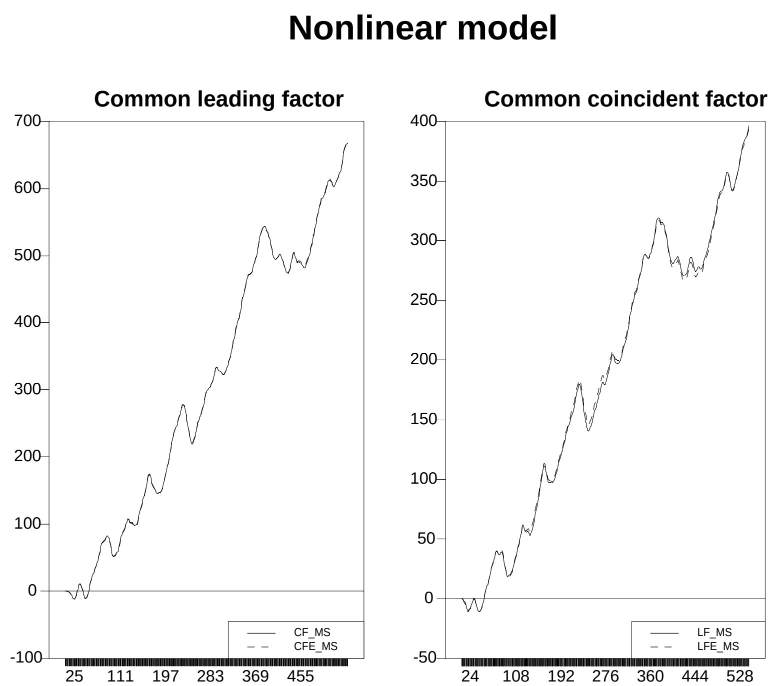

Figure 2: True and estimated common factors

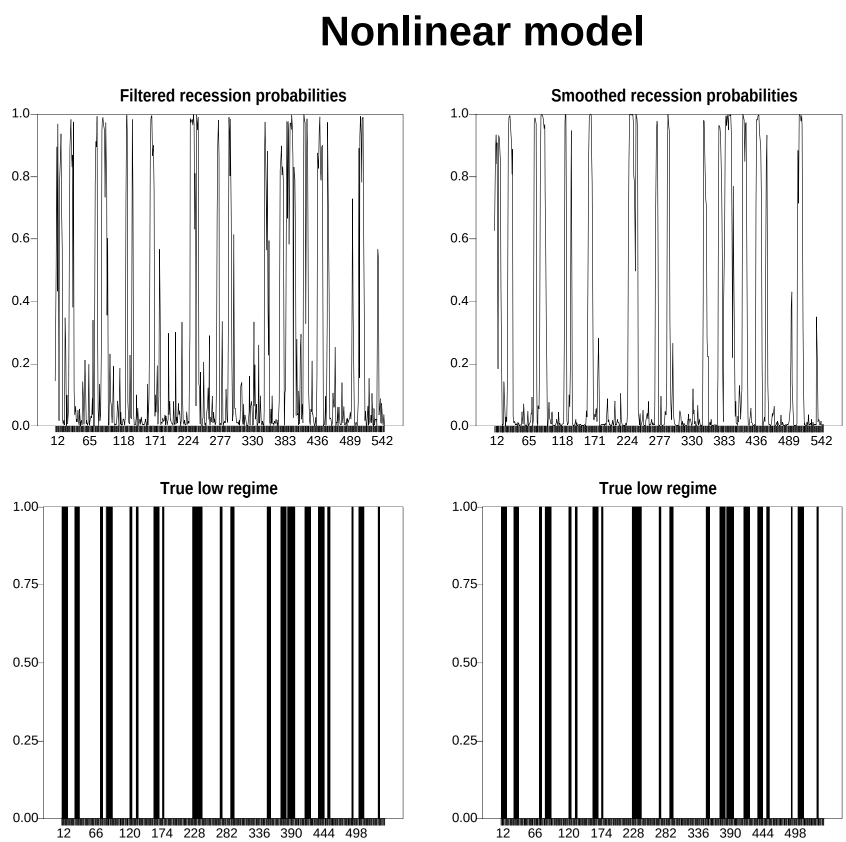

Figure 3: True low regime vs. filtered and smoothed recession probabilities